BIOINFORMATICS SEMINAR

NCBI's Conserved Domain Database

Stephen H. Bryant

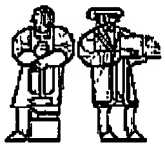
National Center for Biotechnology Information (NCBI), Bldg. 38A Rm. 5S504, National Institutes of Health, 8600 Rockville Pike, Bethesda MD 20894 USA.

Abstract:

Protein evolution has employed a repertoire of a few thousand elementary modules or domains, which form the building blocks of today's proteins. Since structure and molecular function is largely conserved within domain families, computational methods for domain identification have become powerful tools in structure-function analysis. In the talk I will describe the domain identification resources currently provided and under development at NCBI. A principle guiding the project is to provide straightforward access for biologists. To this end conserved-domain alignments from several sources are now included in the PubMed retrieval system, with links to and from the proteins containing domains of that type. Links are computed using the reverse PSI-BLAST algorithm and updated as new structures, sequences, and domain alignments become available. Results are presented as both simple summaries and detailed alignments linked to molecular graphics. Another principle guiding the project is to use protein 3D-structure and the diversity of sequence information now available to improve the accuracy of domain alignments and their utility as reagents for structure-function analysis. To this end curators at NCBI construct sequence-structure alignments representing the conserved three-dimensional cores of domain families, with explicit annotation of conserved binding or catalytic sites observed in structural complexes. Ancient gene duplications which appear to have produced conserved subfamilies with distinct functions are explicitly annotated via subfamily-specific alignments. By providing simple access and using structure and subfamily hierarchy to improve alignment accuracy and annotation, NCBI's conserved domain database project aims to make the results of genome project and the structural genomics initiative as useful as possible to biologists.
For further information see http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml.

Monday March 3, 2003
11:00 a.m. – 1:00 p.m.
(Talk starts at 11:30)
Building NE43, Room 941

Refreshments at 11am in NE43-941
(LCS, 200 Tech Square, Cambridge, MA)



Massachusetts Institute of Technology
Department of Mathematics &
Theory of Computation Group
Lab for Computer Science
Cambridge, MA 02139

http://www-math.mit.edu/compbiosem/

*: Support for the invitation of Dr. Stephen Bryant (NCBI) and of Dr. Juergen Stelling (Magdeburg) is generously provided by ***SERONO Reproductive Biology Institute in Boston (SRBI)*** , www.serono.com with special thanks to Dr. Jadwiga Bienkowska.